植 物 分 类 与 资 源 学 报 2014, **36** (3): 331~341
Plant Diversity and Resources DOI: 10.7677/ynzwyj201413123

MIR166 基因家族在陆生植物中的进化模式分析*

赵旭耀[1,2], 陈斯云[1], 赵 磊[1], 张雪梅[1], 马朋飞[1], 郭振华[1**]
(1 中国科学院昆明植物研究所中国西南野生生物种质资源库, 云南 昆明 650201; 2 中国科学院大学, 北京 100049)

摘要: MicroRNA (miRNA) 是一类广泛存在于真核生物中的具有转录后水平调控功能的内源非编码小分子 RNA。在植物中, miRNA 通过对靶基因的剪切或沉默来实现对植物生命活动的调控, 它是基因表达调控网络的重要组成部分。miR165/166 (miR166) 是陆生植物中最为古老的 *MIRNA* 家族之一, 它通过对 3 型同源异域型-亮氨酸拉链 (*HD-ZIP III*) 等靶标的调控, 在植物的众多发育时期起着关键的调控作用。本文分析了 *MIR166* 基因在陆生植物中的进化关系, 并对 *MIR166* 在基部陆生植物小立碗藓 (*Physcomitrella patens*) 中的复制及进化进行了研究。此外, HD-ZIP III 蛋白是植物中重要的一类转录因子, miR166 对 *HD-ZIP III* 基因的调控作用在陆地植物保守的存在, 本文对 *HD-ZIP III* 基因和 miR166 在进化中的相互作用进行了初步的探讨。

关键词: miR166; *MIRNA* 基因进化; 基因重复; HD-ZIP III 转录因子
中图分类号: Q 75 **文献标识码**: A **文章编号**: 2095-0845(2014)03-331-11

Evolution of *MIR166* Gene Family in Land Plants

ZHAO Xu-Yao[1,2], CHEN Si-Yun[1], ZHAO Lei[1], ZHANG Xue-Mei[1], MA Peng-Fei[1], GUO Zhen-Hua[1**]
(1 *Germplasm Bank of Wild Species in Southwest China*, *Kunming Institute of Botany*, *Chinese Academy of Sciences*, Kunming 650201, China; 2 *University of Chinese Academy of Sciences*, Beijing 100049, China)

Abstract: MicroRNA (miRNA) is a class of endogenous non-coding small RNAs with important post-transcriptional regulatory roles in eukaryotes. Plant miRNAs play important roles in the post-transcriptional regulatory network through mediating cleavage or silencing of target mRNAs. As one of the ancient *MIRNA* family, miR165/166 (miR166) is a key regulator in land plants. In this study, we analyzed the diversity and molecular evolution of *MIR166* genes in land plants, and identified the replication and evolution of *MIR166* genes in *Physcomitrella patens*. Homeodomain Leucine-zipper of class III (HD-ZIP III) proteins are important transcription factors in plants, *HD-ZIP III* genes are conserved targeting of miR166 in angiosperms, gymnosperms, ferns, and mosses. Here, a preliminary study was conducted to make clear the relationships between *HD-ZIP III* genes and miR166 during the evolutionary process.

Key words: miR166; *MIRNA* gene evolution; gene duplication; HD-ZIP III transcription factors

After the first miRNA gene (*lin-4*) was identified in *Caenorhabditis elegans* (Lee *et al.*, 1993), more and more miRNAs were discovered in eukaryotes through deep sequencing and bioinformatics

* Funding: The NSFC-Yunnan province joint foundation (U1136603); the Knowledge Innovation Project of the Chinese Academy of Sciences (KSCX2-YW-N-067); the National Natural Science Foundation of China (30990244); Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry and the Young Academic and Technical Leader Raising Foundation of Yunnan Province (2008PY065); the Western Light Talent Culture Project of the Chinese Academy of Sciences (2010312D11035); and the Yunnan Provincial Government through an innovation team programme

** Author for correspondence; E-mail: guozhenhua@mail.kib.ac.cn
Received date: 2013-05-28, Accepted date: 2014-01-03
作者简介: 赵旭耀 (1987-) 男, 硕士, 从事生物信息与比较基因组学研究。E-mail: zhaoxuyao@mail.kib.ac.cn

approaches. Those *MIRNA* genes with significant sequence homology to each other when compared based on mature miRNA sequences are grouped into miRNA families (Ambros *et al.*, 2003). Hundreds of miRNA families have been found in plants, but most of them are lineage-specific, suggesting that most known *MIRNA* genes arose relatively recently in evolutionary time scale (Cuperus *et al.*, 2011). There are only 8 families are deep conserved in land plants, miR166 family is one of this kind (Nozawa *et al.*, 2010).

miRNAs play important roles in various developmental and physiological processes (Seed germination, vegetative growing, flowering, seeding and resistance to various abiotic or biotic stress) by targeting mRNAs for cleavage or translational repression at the post-transcriptional level in the plants kingdoms (Reinhart *et al.*, 2002; Carrington and Ambros, 2003). miR166 and its targets regulate an array of plant developmental processes, including shoot apical and lateral meristem formation, leaf polarity, floral development, and vascular development (McConnell *et al.*, 2001; Otsuga *et al.*, 2001; Prigge *et al.*, 2005; Jung and Park, 2007).

An increasing number of plant miRNA genes have been discovered through experimental and bioinformatic studies in plants, and they provide an opportunity to study the origins and evolution of them. There have been some evolutionary research on miRNAs in animals, plants and some smaller lineage (Li *et al.*, 2011; Hertel *et al.*, 2012; Zhao *et al.*, 2012). The miRNAs in land plants are regarded as deriving from a common ancestor (Jones-Rhoades, 2012). However, it is unclear whether the miRNA pathway in green alga shares a common ancestry with miRNA pathways of land plants, or whether it is independently derived. There have not found a miRNA conserved between land plants and algas (Jones-Rhoades, 2012).

The previous research had revealed the evolutionary history of the *MIR166* homologs in *Arabidopsis* derived from duplication events, including tandem duplications and segmental duplication events (Maher *et al.*, 2006). Then, Sun *et al.* (2012) characterized the *MIR166* genes derived from duplication events in rice, sorghum, and poplar. Those researches indicating that duplication events are important in the expansion of *MIR166* genes in flowering plants. In this study, we found the phenomenon is exist in moss too, then, the overall phylogenetic tree of miR166 in *Physcomitrella patens* was reconstructed base on the duplication events.

In addition to the duplication and loss events in the genome, the targets of miRNA is another important factor in the miRNA evolution. The "new" miRNAs only have a few suspected targets while the conserved miRNAs target numerous important mRNAs in the eukaryotes. In the evolutionary history of miRNAs, the "ancestor" miRNAs were subdivided into hundreds of families, and the miRNAs which capture important targets finally evolve to the conserved multi-gene families (Bompfunewerer *et al.*, 2005; Hertel *et al.*, 2012). The mature miRNA sequence is complementary to the mRNA targeted region, thus the conserved target determine the conservative property of miRNA. The gain of target mRNA can cause the expansione or functional diversification of miRNA, and the loss and pseudogene change of mRNA can cause deletion or functional loss of miRNA.

In this study, we reconstructed the phylogenetic tree of *MIR166* genes in land plants, then, we identified the duplication events originated miR166 genes in *Physcomitrella patens* and inferred their phylogenetic relationship. At last, the phylogenetic distribution of miR166 targeted- HD-ZIP III in land plants were constructed to study the interaction between miR166 and HD-ZIP III.

1 Materials and methods

1.1 Phylogenetic analysis of MIR166 genes in land plants

The mature sequences and precursor sequences of miRNAs were downloaded from miRBase database release 19 (http://www.mirbase.org/) (Addition-

al file 1, http://journal.kib.ac.cn/UserFiles/File/ZXY.rar). The LOGO representation of mature miRNAs was obtained with the WebLogo software (Crooks *et al.*, 2004).

All of the stem-loop sequences of the miR166 family were aligned using Clustal W (Larkin *et al.*, 2007). Then, We excluded the incorrect aligned sequences in the conserved regions of miRNA: miRNA* and aligned the remaining sequences again. Secondary structures were produced for each sequence in the alignment using RNAfold (Mohsen *et al.*, 2009). Then, we used 4SALE to correct the alignment manually by considering the agreement in secondary structure (Seibel *et al.*, 2006) and the consensus sequences and structure using RNAalifold (Bernhart *et al.*, 2008).

Neighbor-joining (NJ) phylogenetic trees based on the p-distance and Kimura 2-parameter distance were generated by MEGA version 5.0 (Tamura *et al.*, 2011). Bootstrap confidence values were obtained applying by 1 000 replications, and only the clades with bootstrap value higher than 50 were shown.

1.2 Identification of miR166 homologues resides in duplicated blocks in *Physcomitrella patens*

In order to determine whether the miRNAs arose or evolved from segmental duplication events, genome-wide analysis was undertaken to examine whether a miRNA resides within a duplicated block as previously described (Zhang *et al.*, 2009; Sun *et al.*, 2012). First, we used NCBI map viewer (http://www.ncbi.nlm.nih.gov/projects/mapview/) to extract 10 protein-coding genes in flanking regions for each *MIR166* homologues or tandemly duplicated miR166 clusters. Then the flanking protein-coding genes of every miRNA were aligned, using standalone BLAST (blastn, version 2.2.27), against those protein-coding genes surrounding another miRNA in order to identify paralogs. For each miRNA pair, the number of protein coding genes with the best non-self match to protein-coding genes flanking another miRNA was counted (Sun *et al.*, 2012) (Additional file 2, http://journal.kib.ac.cn/UserFiles/File/ZXY.rar). The flanking noncoding sequences were aligned by Emboss to help resolve the evolutionary history of the miR166 family (Rice *et al.*, 2000).

1.3 miR166 targeted-*HD-Zip III* in land plants

A total of 60 mRNA sequences of HD-Zip III protein were downloaded from NCBI. Then 36 sequences from 9 plant species (*Arabidopsis thaliana*, *Medicago truncatula*, *Oryza sativa*, *Pinus taeda*, *Populus trichocarpa*, *Ricinus communis*, *Selaginella moellendorffii*, *Solanum lycopersicum*, and *Sorghum bicolor*) were collected for target prediction (Additional file 3, http://journal.kib.ac.cn/UserFiles/File/ZXY.rar). The miR166-targeted HD-Zip III mRNAs were predicted in psRNATarget under the default rules (Dai and Zhao, 2011).

1.4 Phylogenetic analysis of HD-Zip III transcription factor

Sequences of HD-Zip III mRNA were aligned using CLUSTALX, then MEGA version 5.0 was used to construct neighbor-joining (NJ) and maximum likelihood (ML) tree. Supporting values were assessed using 1000 replicate bootstrap tests.

2 Results

2.1 *MIR166* is a diversity ancient miRNA gene family that is ubiquitously distributed in flowering plants, gymnosperms, ferns, and mosses

So far, miRBase release 19 has collected a total of 209 *MIR166* genes of 35 species ranged from moss to the higher plants (Table 1). The precursor sequences of miR166 are diversified in land plants, the longest precursor is 670nt and the shortest one is 72nt (pvu-miR166 and sbi-miR166b). However, the short functional mature sequence is highly conserved in land plants. We also found that the mature miRNA were sequence characterized in the different lineages (Figure 1). *MIR166* is a typical multiple-gene family, and there are numerous paralogs in many plants species (miR166 is single copy in 7 plants, but it's attribute to the lack of genome and deeper research). Evidences above indicate that *MIR166* is an old miRNA family in land plants. The

features of miR166 precursors suggest that the *MIR166* genes have a complex evolution history.

Table 1 The distribution of miR166 in land plants

Clade			Species	Number of miR166 genes
Bryophyta			*Physcomitrella patens*	13(13,0)
Fern			*Selaginella moellendorffii*	3(1,2)
Seed plants	Gymnnos-perms		*Picea abies*	2(2,0)
			Pinus densata	2(2,0)
			Pinus taeda	2(2,0)
	Flowering plants	Dicots	*Brachypodium distachyon*	7(5,2)
			Hordeum vulgare	3(2,1)
			Oryza sativa	14(8,6)
			Sorghum bicolor	11(8,3)
			Zea mays	14(12,2)
		Monocots	*Arabidopsis lyrata*	9(4,5)
			Aquilegia coerulea	5(3,2)
			Arabidopsis thaliana	9(4,5)
			Brassica napus	6(4,2)
			Cucumis melo	9(6,3)
			Citrus reticulata	2(2,0)
			Citrus sinensis	5(3,1)*
			Citrus trifoliata	1(0,0)*
			Digitalis purpurea	2(1,1)
			Gossypium hirsutum	1(1,0)
			Glycine max	21(6,15)
			Hevea brasiliensis	2(2,0)
			Helianthus paradoxus	1(1,0)
			Helianthus petiolaris	1(0,1)
			Malus domestica	9(4,5)
			Manihot esculenta	1(0,1)
			Medicago truncatula	8(6,2)
			Nicotiana tabacum	8(3,5)
			Populus trichocarpa	17(9,8)
			Phaseolus vulgaris	1(0,0)*
			Ricinus communis	5(4,1)
			Solanum lycopersicum	2(2,0)
			Salvia sclarea	1(1,0)
			Theobroma cacao	4(4,0)
			Vitis vinifera	8(2,6)

The number in the bracket is type 1 and type 2 miR166 gene in this species, * represent the miRNA genes with 5' mature miRNAs

2.2 Evolution of *MIR166* gene in land plants

Stem-loop sequences were usually applied to study the phylogenetic relationship because they are the most conserved parts of the miRNA genes (Li *et al.*, 2011; Hertel *et al.*, 2012; Zhao *et al.*, 2012). The selections for mutations in structured RNA mainly arise from base-pairings to maintain the RNA structures while in protein-coding sequences selection pressure disadvantages mutations in triplet codes that disrupt protein functions through alteration of amino-acids (Li *et al.*, 2011). As a class of functional RNA, structure character of pre-miRNA is more important than its sequence character out of the region of miRNA:miRNA*, just as nonsense mutation in the protein, the substitution don't cause structural change in the stem-loop is non-signification mutation. So we construct the phylogenetic tree of *MIR166* using the structure and sequence characters of the stem-loop.

We aligned the stem-loop sequences of miR166, three precursors with 5' mature miRNA were excluded (Ambros *et al.*, 2003). Then we used RNAfold to predicted their stem-loop structures, in this step, we classified the pre-miR166 into two types according to their structural conservation in the region of loop (Figure 2), type 1 pre-miR166 which have a long stem with several mismatch bases and a loop is canonical stem-loop structure, and there are extra loops in type 2 pre-miR166 cause by the abundant mismatch in this region which may be induced by nucleotide variation in the evolution history. Type 1 is the main form of pre-miR166 in land plants, all the pre-miR166 in the base of land plants *Physcomitrella patens* are belong to type 1, and type 2 pre-miR166 is a variation of type 1.

So we determined to use the 127 type 1 pre-miR166 to build the phylogenetic tree in land plants, the stem-loop are divided into three space partitions in the structure: miRNA:MIRNA* duplexes, prolonged stem, and loop. The most conserved miRNA:MIRNA* duplexes region were retain, in the prolonged stems and loops which is cause by base pairing and base mismatch respectively, the sequence characters of AUCG were translated into structure characters. Then, we used RNAalifold to identify the consensus sequences and structures of prolonged stems and loops of pre-mir166 (Figure 3). At last, the structure

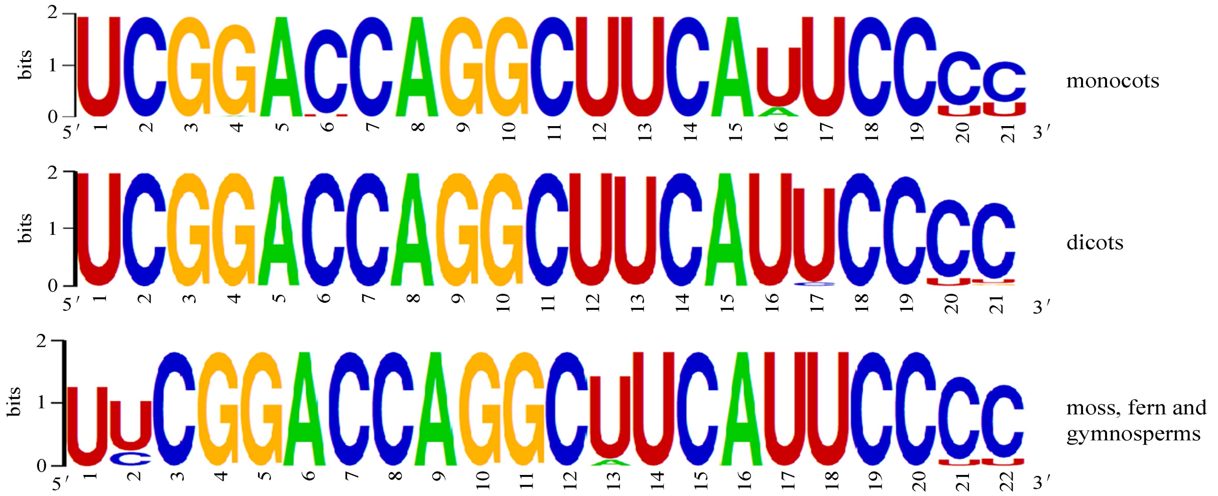


Fig. 1　RNA logo of mature miR166 in land plants

Sequence characterized of the mature miRNA in the different lineages, U/A at16nt in monocots, U/C at 17nt in dicots, U/A at 13nt in the lineage of moss, fern and gymnosperms

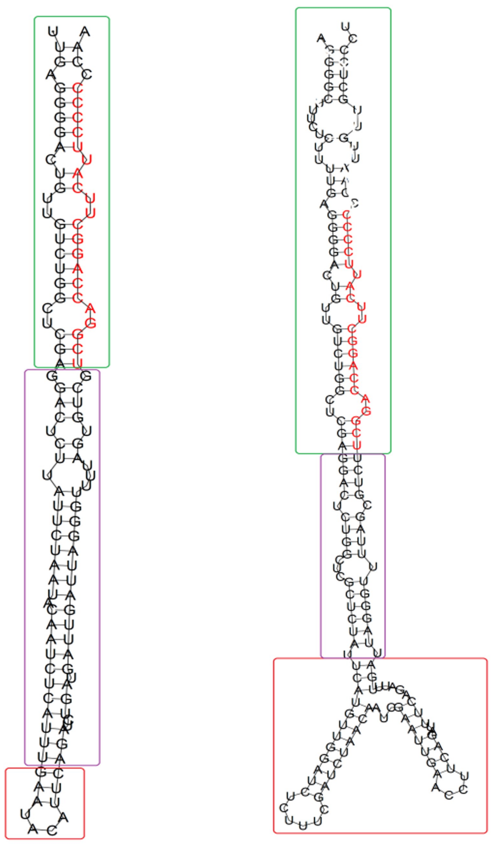


Fig. 2　Structure of pre-miRNA

The stem-loop structure of pre-miRNA is divided into three parts, the part in green wireframe is miRNA: miRNA* duplex, the part in purple wireframe is prolonged stem, the part in red wireframe is loop. The sequences of red is mature miRNA. Type 1 premiRNA in the left have a typical stem-loop secondary structure, and the premiRNA with a complicated loop in the right is belongs to type 2

characters of prolonged stems and loops were translated into sequence characters based on the consensus sequences and structures again (Additional file 4, http://journal.kib.ac.cn/UserFiles/File/ZXY.rar), the sequences of pre-miRNAs were aligned by Clustal W and then the NJ tree were constructed by MEGA 5.0 (Figure 4).

In the tree, the homologous genes are highly conserved in moss, most of *MIR166* genes show a clustered tendency in monocots, they also have a closer neighbor relationship to the moss. However, their homologies in dicots showed a more diversified evolutionary relationship. We speculate that the *MIR166* gene was originated in the original group land plant just as moss. Then, the original seeding plant inherited the gene and then hand down it to the original Gymnosperm and flowering plants in the origination and evolution of flowering and seeding plants, the ancestor of monocots should be appeared before the dicots even the gymnosperm.

2.3　Evolutionary conservation analysis of miR166 homologues derived from duplication events in *P. patens*

Gene duplications are derived from the duplication events which including whole-genome duplication (WGD), segmental duplications, and local du-



Fig. 3 The methods of construct phylogenetic tree of miRNA gene

The part in red is mature and star miRNA sequences, the part in green is the sequences of loop structure, the part in black is the sequences of prolonged stem. In step 1, the most conserved miRNA: MIRNA* duplexes region is removed, the sequences of prolonged stems and loops are retained. Then, the sequence characters of AUCG were translated into structure characters. After that, the consensus sequences and structures of prolonged stems and loops are identify. Next, the structure characters of prolonged stems and loops were translated into sequence characters based on the consensus sequences and structures again. At last, the removed sequences of miRNA and miRNA* are restored

plications that involve one or two genes known as tandem duplications (Maher *et al.*, 2006). The duplication events also contribute to expansion of miRNA families, especially in the WGD, some protein-coding genes and TE will generate new miRNA genes (Hertel *et al.*, 2006; Maher *et al.*, 2006; Yuan *et al.*, 2011). Previous studies showed that many plant *MIRNAs* are resided in the genomic duplication blocks, the context sequencing of them and them duplication is conserved to their repetition (Sun *et al.*, 2012). So the flanking sequences, including conserved coding genes and conserved non-coding regions flanking can be used for identification the *MIRNA* gene in duplicated blocks.

Maher *et al.* (2006) have studied the evolution of duplication derived *MIRNA165/166* in *A. thaliana*, in their research, the overall evolution of seven miR166 members were reconstructed. Recent research suggests ath-miR165a is origin from ath-miR166b before the origin of ath-miR166a, then a base substitution is introduced into ath-miR165a, in the recent large large-scale duplication event, ath-miR165b and ath-miR166a are originated from the duplicated blocks of ath-miR165a and ath-miR166b respectively. It's in accordance with the results in phylogenetic research above, ath-miR165a and ath-miR166b belong to type 1 *MIR166* gene, ath-miR165b and ath-miR166b belong to type 2 *MIR166* gene, it's also a evidence for that the type 2 *MIR166* gene is originated from type 1 *MIR166* gene.

As a genome sequenced plant, the research of miRNA in *P. patens* is concerned. Though *P. patens* is a non-flowering plant, several conserved miRNAs exist in flowering plants were detected in moss. miR166 is the one of the miRNAs which is found in moss earliest, Floyd and Bowman (2004) cloned a *HD-ZIP III* gene homolog *PpC3HDZIP1* which contains a conserved miR166-targerted region from *P. patens*, they also cloned a putative miR166-guided cleavage product of PpC3HDZIP1 mRNA. The experiment indirectly demonstrates the exists of miR166 in moss. In the next study, more and more moss miRNA were detected with the availability of *P. patens* genomes, including 12 *MIRNA166* genes (Arazi *et al.*, 2005). The numerous homologies in moss suggest that the expansion of miR166 is considerable. However, as outlined above, miR166 family is more conserved in *P. patens* than other plants with

many members, so we want to reveal the intraspecific evolution of miR166 in *P. patens*.

In this study, we analysed the context of *MIR166* genes in *P. patens* to seek their flanking coding genes and noncoding sequences, those sequences were used to identify miR166 homology genes origined from duplication block. Then, a total of 9 duplication blocks around *MIR166* genes were obtained in *P. patens* (Table 2). There are 8 out of 13 miR166 loci were located in miRNA clusters (ppt-miR166cdf cluster, ppt-miR166ij cluster, and ppt-miR166hkm cluster), the short length of three miRNA clusters indicate at last one tachytelic evolution events of tandem duplications occurred in moss genome, ppt-miR166m of the base substitution is origin from ppt-miR166h or ppt-miR166k in a latter tandem duplications event. On the other hand, the clusters of miR166 suggest they provide stable and conservative regulatory effect in the *P. patens*. Then through the conserved flanking coding gene between *MIRNA166* genes, we found that one flanking coding gene of miR166e is conserved to two flanking coding genes of the miR166ij cluster, then, the higher conserved stem-loop structure between miR166e and miR166j indicted that miR166j is origin from the duplication block of miR166e before the tandem duplications event. There are two conserved flanking coding genes between ppt-miR166ij and ppt-miR166a, then, we found there are conserved noncoding sequences between ppt-miR166a and ppt-miR166ij, so we inferred that ppt-miR166a is a segmental duplication of ppt-miR166ij cluster before the loss event near the ppt-miR166a.

We also reconstruct the phylogenetic tree of miR166 gene in the *P. patens* (Figure 5). In the tree, we found that the tandem duplications and segmental duplications is frequently in the moss genome, it is most likely to occur a large large-scale duplication event in the genome level, but we need more evidence. The gene loss of miR166 in the moss is frequent too, there are at least 3 loss events we known. There is a replication block between ppt-miR166g

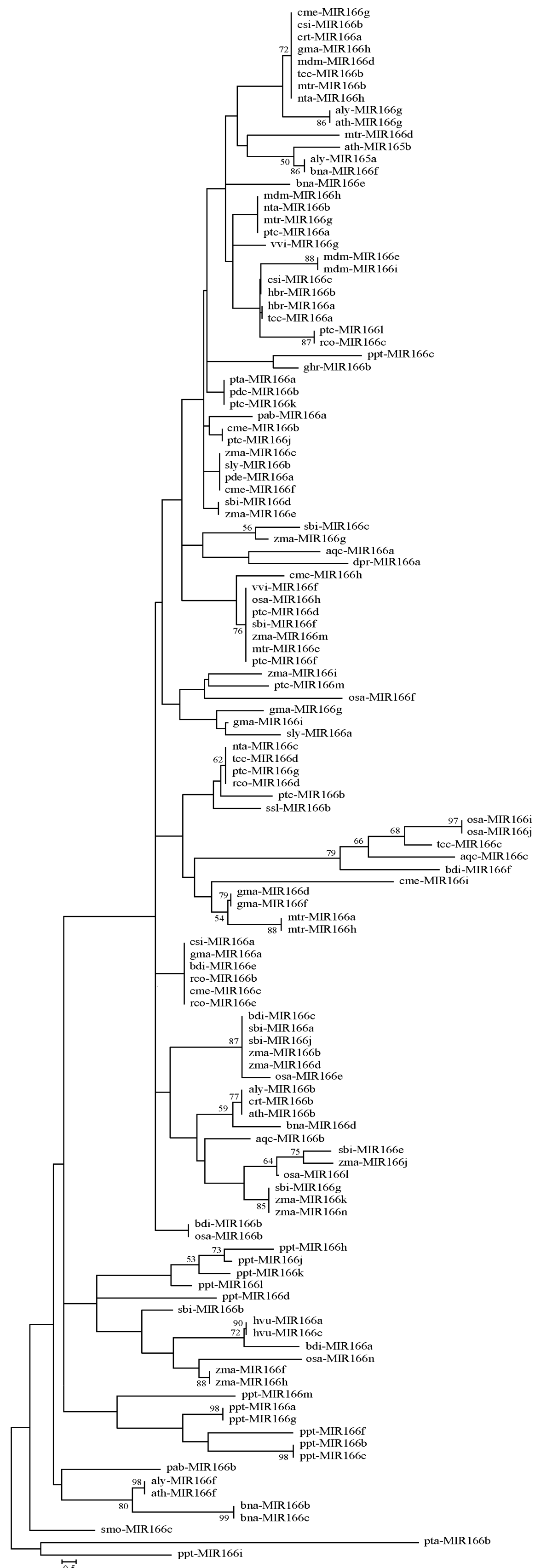


Fig. 4　The phylogenetic tree of type 1 premiR166

Table 2 The number of conserved protein-coding genes in the flanking the miR166

Duplicated miRNAs 1	Duplicated miRNAs 2	flanking conserved protein-coding genes
ppt-miR166a	ppt-miR166ij	2
ppt-miR166b	ppt-miR166e	1
ppt-miR166b	ppt-miR166g	1
ppt-miR166g	ppt-miR166ij	1
ppt-miR166cdf	ppt-miR166l	2
ppt-miR166e	ppt-miR166ij	2
ppt-miR166e	ppt-miR166l	1
ppt-miR166e	ppt-miR166g	1
ppt-miR166ij	ppt-miR166hkm	1

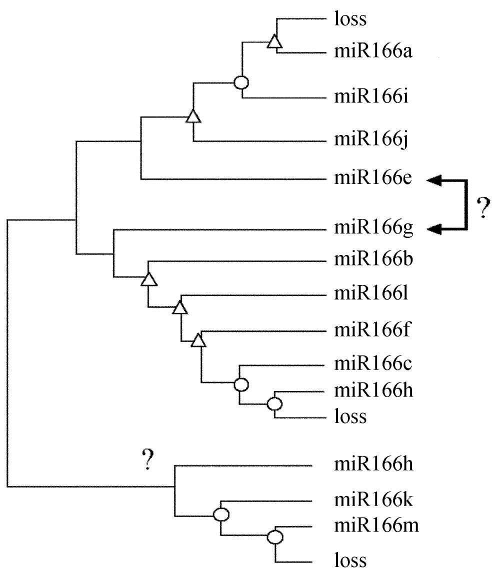


Fig. 5 Reconstruction of miR166 evolution in *Physcomitrella patens* Trigon is represent a segmental duplications event, a circle is a tandem duplications event

and miR166e, there also a replication block of ppt-miR166h between ppt-miR-166i or ppt-miR166l. However, the evolution details of ppt-miR166g and ppt-miR166h is still unknown.

2.4 Phylogenetic distribution of miR166 targeted-*HD-ZIP III* in land plants

Members of the Class III HOMEODOMAIN-LEUCINEZIPPER transcriptional factors are plant-specific and involved in many plant development processes (Boualem *et al.*, 2008; Zhu *et al.*, 2011; Sakaguchi and Watanabe, 2012; Zhang *et al.*, 2012). The shoot apical meristem (SAM) fate is specified by *HD-ZIP III* transcriptional factors (Sakaguchi and Watanabe, 2012). *HD-ZIP III* gene is regulated by miR166: AGO1, in this process, AGO10 combined with miR166 competitively to restraint the negative regulation of *HD-ZIP III* gene and maintenance of undifferentiated stem cells in the SAM (Liu *et al.*, 2009). *HD-ZIP III* gene and *MIR166 gene* are present in many land plants because they make an important contribution to maintain the SAM and determine adaxial/abaxial polarity in plant tissues. Negative regulation of *HD-ZIP III* by miR166 is highly conserved in land plants, and these *HD-ZIP III* functions show diversity among different plant groups (Sakaguchi and Watanabe, 2012).

Twenty-two out of 35 *HD-Zip III* genes were the putative target genes for miR166 family with high probability (Additional file 5, http://journal.kib.ac.cn/UserFiles/File/ZXY.rar). The targeted region of *HD-Zip III* genes is very conserved in land plants. We also built the phylogenetic tree of the eight plants (Figure 6). There are two clades of HD-Zip III genes, all the *P. patens HD-Zip III* genes belong to clade 1, they shows considerable conserved and it may be the reason to the evolutional characters of ppt-miR166. HD-ZIP III protein is a key transcription factor, it seems to be conserved in the duplication events, the more copy of *HD-ZIP III* is a factor of expand of *MIR166* gene. And the stability function of HD-ZIP III restrict the conserved of *MIR166* gene.

3 Discussion

The evolutionary rate around the plant is disproportion, it's more slower in the functional sequences just as protein coding genes than the junk DNA as pseudogene. *MIRNA* gene is a "special case", the overall evolutionary rate of MIRNA gene is approximate to pseudogene although it is transcribed into mRNA-like RNA and generate a functional RNA production finally. The evolutionary rate of the different regions in MIRNA gene is vary drastically, the rate of mature and star miRNA is considerably lower than other regions, the conservative property of mature sequences is higher than the miRNA* sequence

for the mature miRNA are essential to recognize target mRNAs as well as to make a duplex structure with star sequences, whereas star sequences appear to be essential only for keeping the duplex structure with mature sequences (Nozawa *et al.*, 2010).

Individual miRNAs are much more evolutionary labile, the expansion or loss is frequently, then, the incongruous conservative property between mature sequence and precursor sequence increase the difficulty to understand the overall evolutionary of those miRNA gene as *MIR166* through the traditional methods. The structure character is useful to the evolution research in the MIRNA genes.

However, the surrounding conserved coding genes sequences of *MIRNA* gene with normal evolutionary rates, the duplication events and evolutional of them could be traced. So the evolution of the conserved flanking protein-coding gene and non-coding sequences give a track to the evolutional history of miRNAs in the same duplication blocks. The arising of their homologues gene always be associated to the duplications events in the genome, it can contribute

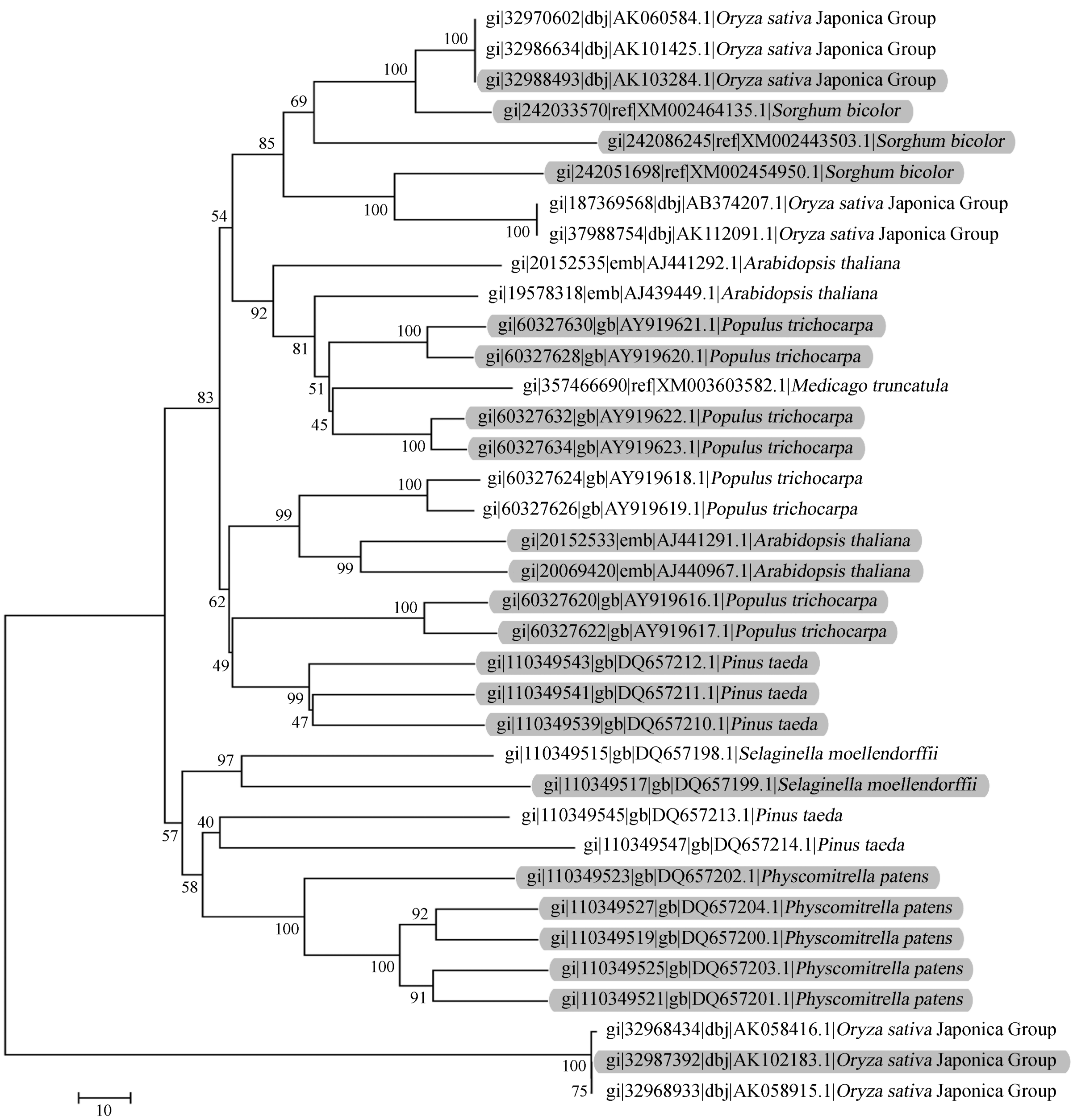


Fig. 6　Phylogenetic distribution of *HD-ZIP III* genes in land plants. The miR166 targeted- *HD-ZIP III* genes are signed in yellow

to the research of the evolution in a smaller lineage.

In reconstructing of the evolution history of *MIR166* genes, the structure characteristics of pre-miR166 were taken into account tentatively and we acquired believable results. We hope miRNAs could be used as the phylogeny markers in the future.

Acknowledgments: We thank Dr. Wei Jiang (Kunming Institute of Botany, Chinese Academy of Sciences) and Pro. Berezikov (Hubrecht Institute, KNAW and University Medical Center Utrecht) for their constructive advices.

References:

Ambros V, Bartel B, Bartel DP *et al*., 2003. A uniform system for microRNA annotation [J]. *RNA* (A Publication of the RNA Society), **9** (3): 277—279

Arazi T, Talmor-Neiman M, Stav R *et al*., 2005. Cloning and characterization of micro-RNAs from moss [J]. *The Plant Journal*, **43** (6): 837—848

Bernhart SH, Hofacker IL, Will S *et al*., 2008. RNAalifold: improved consensus structure prediction for RNA alignments [J]. *BMC Bioinformatics*, **9**: 474—481

Bompfunewerer AF, Flamm C, Fried C *et al*., 2005. Evolutionary patterns of non-coding RNAs [J]. *Theory in Biosciences*, **123** (4): 301—369

Boualem A, Laporte P, Jovanovic M *et al*., 2008. MicroRNA166 controls root and nodule development in Medicago truncatula [J]. *The Plant Journal*, **54** (5): 876—887

Carrington JC, Ambros V, 2003. Role of microRNAs in plant and animal development [J]. *Science*, **301** (5631): 336—338

Crooks GE, Hon G, Chandonia JM *et al*., 2004. WebLogo: A sequence logo generator [J]. *Genome Research*, **14** (6): 1188—1190

Cuperus JT, Fahlgren N, Carrington JC, 2011. Evolution and functional diversification of MIRNA genes [J]. *The Plant Cell*, **23** (2): 431—442

Dai XB, Zhao PX, 2011. psRNATarget: a plant small RNA target analysis server [J]. *Nucleic Acids Research*, **39**: W155—W159

Floyd SK, Bowman JL, 2004. Gene regulation: Ancient microRNA target sequences in plants [J]. *Nature*, **428** (6982): 485—486

Hertel J, Bartschat S, Wintsche A *et al*., 2012. Evolution of the let-7 microRNA Family [J]. *RNA Biology*, **9** (3): 231—241

Hertel J, Lindemeyer M, Missal K *et al*., 2006. The expansion of the metazoan microRNA repertoire [J]. *BMC Genomics*, **7**: 25—39

Jones-Rhoades MW, 2012. Conservation and divergence in plant microRNAs [J]. *Plant Molecular Biology*, **80** (1): 3—16

Jung JH, Park CM, 2007. MIR166/165 genes exhibit dynamic expression patterns in regulating shoot apical meristem and floral development in *Arabidopsis* [J]. *Planta*, **225** (6): 1327—1338

Larkin MA, Blackshields G, Brown NP *et al*., 2007. Clustal W and clustal X version 2.0 [J]. *Bioinformatics*, **23** (21): 2947—2948

Lee RC, Feinbaum RL, Ambros V, 1993. The C-Elegans heterochronic gene Lin-4 encodes small rnas with antisense complementarity to Lin-14 [J]. *Cell*, **75** (5): 843—854

Li Y, Li CQ, Ding GH *et al*., 2011. Evolution of MIR159/319 microRNA genes and their post-transcriptional regulatory link to siRNA pathways [J]. *BMC Evolutionary Biology*, **11**: 122

Liu QL, Yao XZ, Pi LM *et al*., 2009. The ARGONAUTE10 gene modulates shoot apical meristem maintenance and establishment of leaf polarity by repressing miR165/166 in Arabidopsis [J]. *The Plant Journal*, **58** (1): 27—40

Maher C, Stein L, Ware D, 2006. Evolution of *Arabidopsis* microRNA families through duplication events [J]. *Genome Research*, **16** (4): 510—519

McConnell JR, Emery J, Eshed Y *et al*., 2001. Role of PHABULOSA and PHAVOLUTA in determining radial patterning in shoots [J]. *Nature*, **411** (6838): 709—713

Mohsen AM, Khader AT, Ramachandram D, 2009. Comparison of HSRNAFold and RNAFold Algorithms for RNA Secondary Structure Prediction [J]. Tencon 2009-2009 Ieee Region 10 Conference, Vols 1-4, 1557—1562

Nozawa M, Miura S, Nei M, 2010. Origins and evolution of microRNA genes in plants [J]. *Genes & Genetic Systems*, **85** (6): 398—398

Otsuga D, DeGuzman B, Prigge MJ *et al*., 2001. REVOLUTA regulates meristem initiation at lateral positions [J]. *The Plant Journal*, **25** (2): 223—236

Prigge MJ, Otsuga D, Alonso JM *et al*., 2005. Class III homeodomain-leucine zipper gene family members have overlapping, antagonistic, and distinct roles in *Arabidopsis* development [J]. *The Plant Cell*, **17** (1): 61—76

Reinhart BJ, Weinstein EG, Rhoades MW *et al*., 2002. MicroRNAs in plants [J]. *Genes & Development*, **16** (13): 1616—1626

Rice P, Longden I, Bleasby A, 2000. EMBOSS: The European molecular biology open software suite [J]. *Trends in Genetics*, **16** (6): 276—277

Sakaguchi J, Watanabe Y, 2012. miR165/166 and the development of land plants [J]. *Development Growth & Differentiation*, **54** (1): 93—99

Seibel PN, Muller T, Dandekar T *et al*., 2006. 4SALE - A tool for synchronous RNA sequence and secondary structure alignment and editing [J]. *BMC Bioinformatics*, **7**: 498

Sun J, Zhou M, Mao ZT *et al*., 2012. Characterization and evolution of microRNA genes derived from repetitive elements and duplication events in plants [J]. *PLOS ONE*, **7** (4): e34092

Tamura K, Peterson D, Peterson N *et al*., 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods [J]. *Molecular Biology and Evolution*, **28** (10): 2731—2739

Yuan ZD, Sun X, Liu HD *et al*., 2011. MicroRNA genes derived from repetitive elements and expanded by segmental duplication events in mammalian genomes [J]. *PLOS ONE*, **6** (3): e17666

Zhang CH, Zhang YP, Guo L *et al.*, 2012. Characterization of the miR165 family and its target gene Pp-ATHB8 in *Prunus persica* [J]. *Scientia Horticulturae*, **146**: 21—28

Zhang LF, Chia JM, Kumari S *et al.*, 2009. A genome-wide characterization of MicroRNA genes in maize [J]. *PLOS Genetics*, **5** (11): e1000716

Zhao JP, Diao S, Zhang BY *et al.*, 2012. Phylogenetic analysis and molecular evolution patterns in the MIR482-MIR1448 Ppolycistron of *Populus* L [J]. *PLOS ONE*, **7** (10): e47811

Zhu HL, Hu FQ, Wang RH *et al.*, 2011. Arabidopsis argonaute10 specifically sequesters miR166/165 to regulate shoot apical meristem development [J]. *Cell*, **145** (2): 242—256

Additional file

Additional file 1. The sequences and structures of type 1 premiR166.

Additional file 2. The number of conserved protein-coding genes in the flanking the miR166.

Additional file 3. mRNA sequences of HD-ZIP III we used for target prediction.

Additional file 4. The premiR166 sequences processed through the structure character.

Additional file 5. miR165/166 targeted-HD-ZIP III genes in land plants.

〔上接第 **296** 页〕

华南植物研究所早期史——中山大学农林植物研究所史事（1928-1954），中国近世生物学机构与人物丛书，胡宗刚著，上海交通大学出版社出版；2013，216 页。ISBN：978-7-03-039267-1.

中国科学院华南植物园前身是中山大学农林植物研究所，1928 年由著名植物学家陈焕镛在中山大学农科设立。陈焕镛先生以其雄才博学，致力于种类丰富的华南地区植物区系研究，使中大农林植物研究所与胡先骕主持的北平静生生物调查所植物部“双星辉映”，成为民国时期一南一北两个最重要的植物学研究机构。1954 年，中山大学农林植物研究所改隶中国科学院，成为中国科学院华南植物研究所（现名中国科学院华南植物园）。长期以来，学界对中山大学农林植物研究所（即 1954 年改隶中国科学院之前）这段早期历史，未曾有过系统梳理，许多人物和事迹已湮没不彰。本书以档案为主要材料，记述中山大学农林植物所发展脉络，事不分巨细，凡有记录，即采入书中，从细微处，或可见历史真实。同时，陈焕镛为该所奠基人，长期执掌该所，中大植物所发展与陈焕镛个人经历密不可分，故本书也是对陈焕镛学术经历作完整记述。全书分为：引言、第一章：陈焕镛家世与其人、第二章：农林植物所之创建与发展、第三章：播迁与流亡、第四章：战后复员、第五章：改隶中国科学院、大事记和参考文献。

西双版纳热带植物园五十年，胡宗刚著，科学出版社出版；2014，370 页。ISBN：978-7-03-039267-1.

中国科学院西双版纳热带植物园由著名植物学家蔡希陶主持创建，经过五十年发展，已成为中国热带植物学的重要研究机构，也是西双版纳重要旅游景区。本书以该园所藏档案为主要材料，兼而采访相关人士，以科学社会学方法撰写而成，完整再现该园五十年历程。由于机构几经分合改隶，本书还囊括并入机构之始末，如云南热带森林生物地理群落定位研究站、中国科学院昆明生态研究所等，所涉及主要人物除蔡希陶外，还有许再富、裴盛基、冯耀宗、陈进等。书中记录了大量著名学者和知名人士与该园的相关活动，许多史实系首次披露，所附大量照片多为首次公布。全书共分为六章：创建始末、建园初期、文化大革命十年、改革之中求发展、分途发展、重新整合。书末附有结语、编年纪事、主要参考文献、人名索引和后记。

胡宗刚先生不愧为中国植物学史研究之骄子！看过他过去十年发表的一系列有关著作：**不该遗忘的胡先骕**（长江文艺出版社，2005）、**静生生物调查所史稿**（山东教育出版社，2005）、**胡先骕先生年谱长编**（江西教育出版社，2008）、**王文采口述自传**（湖南教育出版社，2008）、**庐山植物园最初三十年**（上海交通大学出版社，2009）、**北平研究院植物学研究所史略**（上海交通大学出版社，2010），真的让人赞叹不已！不要说自学成才了，就是我等专业人士也自愧不如。

马金双（邮箱：jinshuangma@gmail.com）
2014 年 4 月 15 日星期二于上海松江